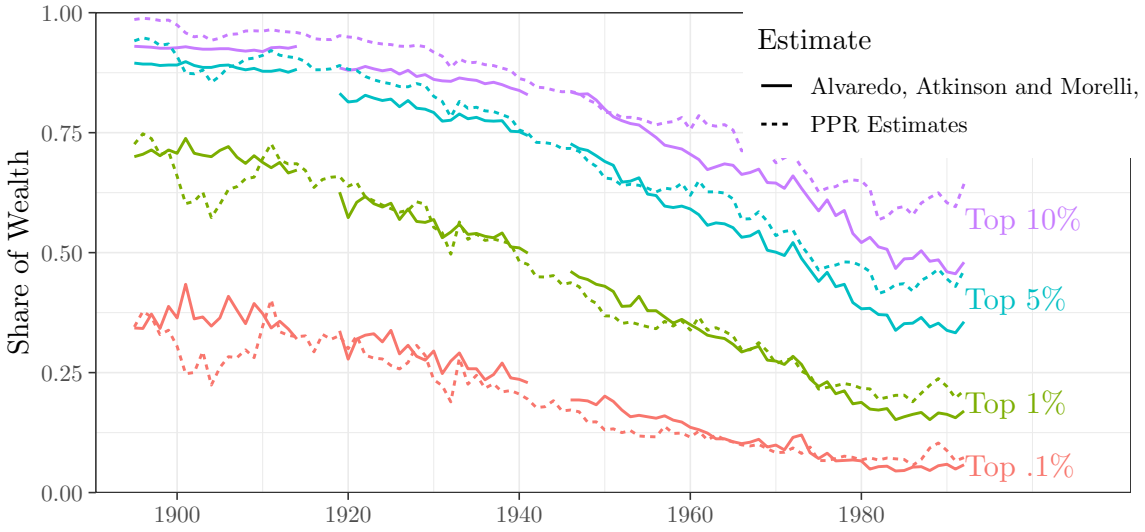
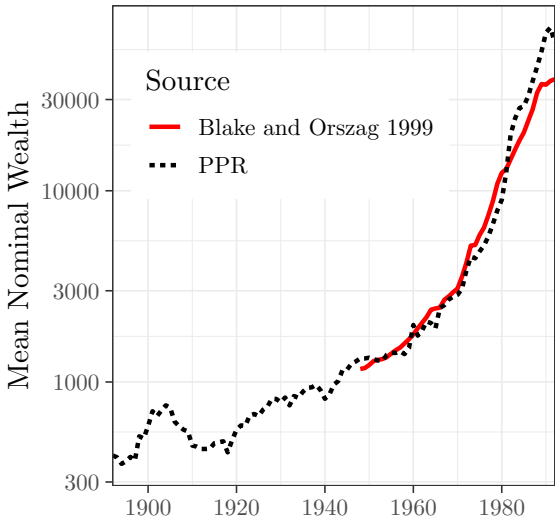
ONLINE APPENDIX to "The Irish in England"

Neil J Cummins and Cormac Ó Gráda

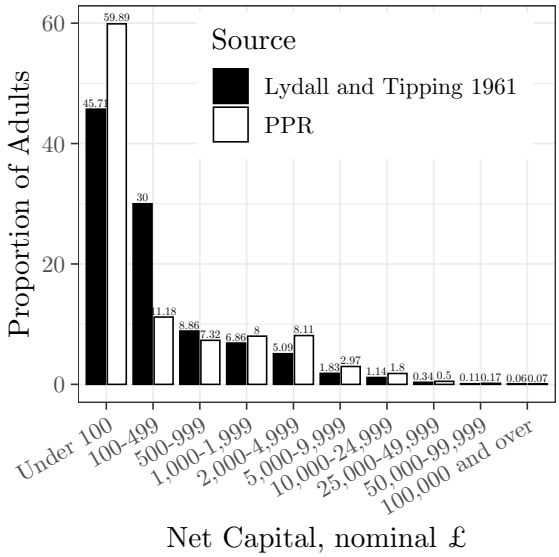
Extra Detail on Data



(a) Comparing Different Estimates of Top Wealth Shares, England 1892-1992



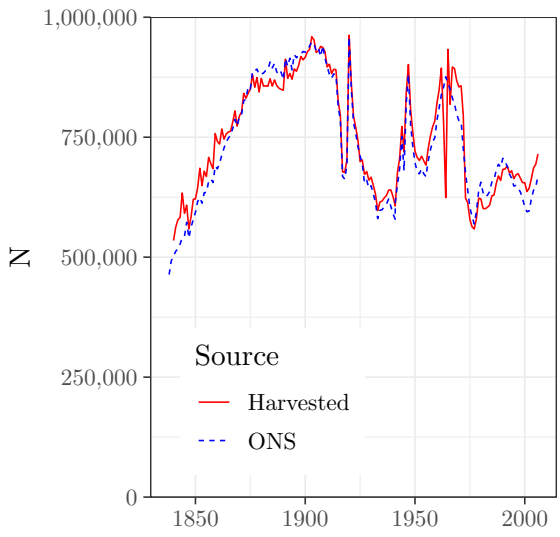
(b) Comparing Average Wealth with Blake and Orszag (1999)



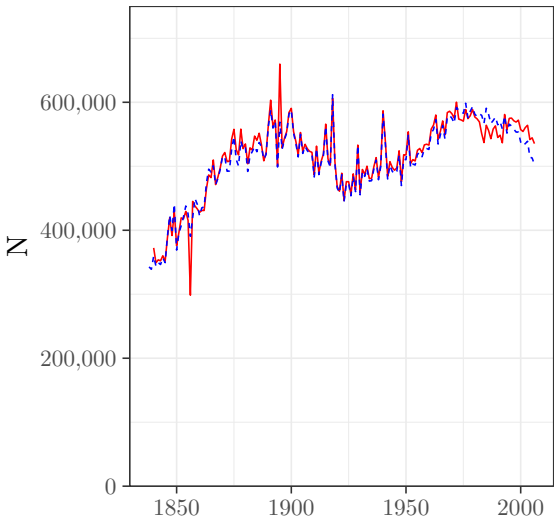
(c) Comparison of Net Capital with Lydall and Tipping (1961), by Wealth Band, 1950s

Figure 12: The PPR Calendar Wealth Data, Compared with Existing Estimates

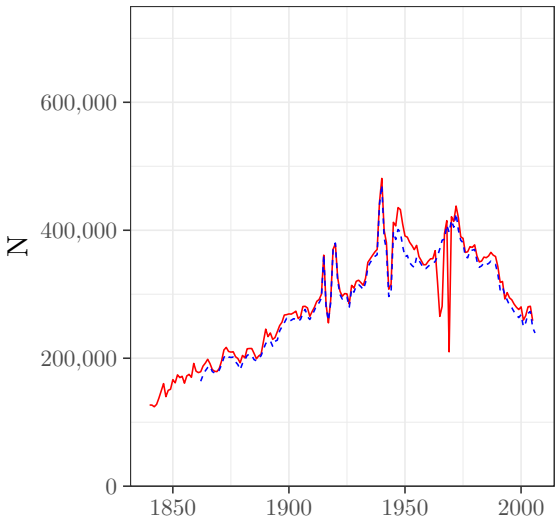
Notes: See Cummins (2021) for a detailed account of the source, construction and validation of the PPR data. Sources: PPR wealth data, Alvaredo et al. (2017) table D1, Blake and Orszag (1999, Table 12) (sum of columns 'net financial wealth', 'housing wealth' and 'consumer durable assets'). These aggregate sums were converted to a per adult measure using population data from Office for National Statistics (2018). Source for figure c: Lydall and Tipping (1961, p.89). Note that the PPR covers England, the Lydall and Tipping (1961) estimates cover Great Britain. Both estimates exclude pension wealth. These figures are also reported in Cummins, 2021.



(a) Births



(b) Deaths



(c) Marriages

Figure 13: Data Collection Verification, 'Harvested' versus Official Count Comparison

Notes: The source for the Offical Counts is Office for National Statistics (2021b).

The Proportion with 'Significant Wealth', 1996-2018

The PPR Calendar data was supplemented by a database of the number of deaths and the number of probates, by surname, 1996-2018. Every probate over this period is listed, by name, on `https://probatesearch.service.gov.uk/#calendar`. It was necessary to enter an exact surname on the webpage to return the count of that surname for a given year. From a 100% sample of the 1881 census ((Schurer and Woollard, 2000)) and the 100% samples of births, marriages and deaths, 1838-2007, and the probate Calendar 1892-1992, a master-list of 3,535,375 surnames was created. Of these surnames many were mistakes so a second list was created filtering the master list by the criteria that the name appeared at least 5 or more times in the death registers, 1983-2007. This resulted in 92,812 surnames which were searched individually for every year 1996 to 2020, a total of 2,320,300 searches for each of the 25 years. (As the probate process can take a few months to a year and those years are this incomplete, we do not report the post 2018 rate here.) Each surname from this master-list was entered into `https://probatesearch.service.gov.uk/#calendar` and the count recorded (GOV.UK, 2018).

As reported in table 7 the threshold estate value above which probate was legally required has been £5,000 from 1984 to today, 2020. In recent years however, the *de facto* reality is that financial institutions have exercised discretion in releasing monies to relatives and beneficiaries from the bank accounts of the recently deceased. In 2020, banks apply their own discretion upon which accounts need probate and which don't. The value they apply as their probate limit could ranges from £5,000 to £50,000.[28]

It is not clear from existing academic literature or the archives of official Govt. websites advising on probate (https://webarchive.nationalarchives.gov.uk) when exactly the nominal probate went from being a flat £5,000 across all institutions, to a discretionary amount that varies in the range £5-£50 thousand, and is institution specific. In 2007-8 (see Atkinson et al. (2017, F8) and as late as 2010 (See Karagiannaki (2015, p.187)), there is evidence that the £5,000 probate threshold was generally applied.[29]

Before 1994, at least, and probably until at least 2010, the assumption that the non-probated estates were worth precisely less than £5,000 appears to be well justified. However, for post-

[28] The current official Government advice on probate states "Contact each asset holder (for example a bank or mortgage company) to find out if you'll need probate to get access to their assets. Every organization has its own rules." GOV.UK (2020). A list of the institution specific probate limits are reported here: https://www.co-oplegalservices.co.uk/media-centre/articles-may-aug-2018/bank-limits-for-probate/. A news article from 1994 states "Although the Act does not specifically apply to banks and to building societies, they usually apply their discretion in a similar way, and will normally only pay out above the pounds 5,000 limit with a grant of probate."' https://www.independent.co.uk/news/business/why-the-bereaved-must-wait-rules-governing-the-release-of-money-when-a-person-dies-can-cause-1420519.html. A 2017 blog post by a probate professional https://www.todayswillsandprobate.co.uk/guest-writers/obtaining-up-to-50k-without-grant-probate/ discusses the change.

[29] Atkinson et al. (2017) state "We have been told by Her Majesty's Revenue & Customs (HMRC) that the 'small estate' category probably accounts for the large majority of estates that do not go through probate " (p.F8).

Years	Nominal Probate Threshold	Source
1858-1900	£10	Turner 2010 p.628
1901-1931	£50	Turner (2010) p.628
1932-1964	£100	Atkinson and Harrison (1978) p.36
1965-1974	£500	Atkinson and Harrison (1978) p.36
1975-1984	£1,500	Atkinson and Harrison (1978) p.36
1984→	£5,000	Turner (2010) p.628, Alvaredo et al. (2018) p.29
		Atkinson et al. (2017) p.F8, Karagiannaki (2015) p.187

Table 7: The Minimum Probate Threshold, 1858-2017

2010, and in particular more recent years, this assumption is not reasonable. Therefore, we can only interpret the probate rate 1996-2018 as being an indicator of wealth that was significant enough for the asset holders (e.g the banks or building societies) to demand an act of probate before transferring the monies. As this could be anywhere between £5,000-50,000, the probate rate after 2010 can only be interpreted as a measure of significant wealth, and not wealth above the legal probate threshold. So I report this measure separately here and do not include it in the main analysis.

Table 8 reports the count of probates, the sum of adult deaths and the proportion probated from 1996 to 2020. The proportion of adult deaths (deaths to those aged 20 and above) requiring an act of probate to deal with their financial assets at death is consistently around 50%. (Note that the 45-47% recorded in 2016-8 may be underestimated due to the lag in recording probates.) This is consistent with the calculations of Karagiannaki (2015) in her analysis of inherited wealth, who estimates a proportion probated of about 50% for the period 2002-2007 (p.187). A figure of 50% is also reported for 2016 in House of Commons Library (2019, p.7).

Table 8: Proportioon Probated, 1996-2018

	$N_{Probates}$	$N_{AdultDeaths}$	Prop. Probated
1996	266,236	556,003	0.48
1997	270,153	551,125	0.49
1998	267,581	546,765	0.49
1999	268,320	546,980	0.49
2000	260,342	531,734	0.49
2001	257,968	526,436	0.49
2002	258,379	529,468	0.49
2003	261,600	533,201	0.49
2004	250,165	508,443	0.49
2005	251,295	507,230	0.50
2006	246,889	496,696	0.50
2007	247,885	498,258	0.50
2008	250,171	503,390	0.50
2009	242,546	485,806	0.50
2010	246,748	488,040	0.51
2011	240,566	479,335	0.50
2012	248,151	494,422	0.50
2013	249,000	502,187	0.50
2014	242,478	496,853	0.49
2015	250,743	525,073	0.48
2016	242,379	520,610	0.47
2017	248,864	528,838	0.47
2018	241,124	537,228	0.45

Source: Office for National Statistics (2019) and `probatesearch.service.gov.uk`

Other Extra Detail

Irish Names

The PPR Calendar data was processed via an OCR (Optical Character Recognition) engine. The process in general worked very well and the resulting data set passed multiple data-quality tests (Cummins (2019)). Amongst the Irish, names beginning with "O'"are commonplace, and non existent within other populations. As the OCR process and the algorithms used to extract surnames may have missed this "", I inspected all possible candidate "O'" stemmed names in

the PPR calendar data. This check turned up numerous oddities. For example, there are 37,613 deaths 1838-2007 for people with the surname "O'Brien" yet only 5 probates recorded, 1858-1992, for this surname. Yet, there are 3,175 probates recorded for the name "Brien" but only 2,304 deaths. I cross-checked all Irish names and assigned any possible stemmed names to the most common occurrence, as measured by the count of all deaths to that name, 1838-2007. Mechanically this was done by summing all deaths in the death data, all probates in the PPR Calendar data and inspecting all 5,805 Irish names for anomalies. This meant that all "Briens" were updated to "O'Brien", "Neill" to "O'Neill", but all "O'Sullivans" were updated to "Sullivan", "O'Daly" to "Daly". Surnames were only updated where both the stemmed and non-stemmed version were of Irish ancestry (thereby grouping over Irish doesn't make any difference to the results).

Surname	N	Updated Surname	N
O'KELLY	448	KELLY	98809
O'SULLIVAN	8085	SULLIVAN	48079
O'RYAN	118	RYAN	40263
BRIEN	2366	O'BRIEN	39108
MCGOUGH	1937	GOUGH	30224
O'CONNOR	21934	CONNOR	28858
MCCARROLL	514	CARROLL	28031
O'CARROLL	433	CARROLL	28031
O'BYRNE	411	BYRNE	25229
MCQUINN	293	QUINN	25153
MCFLYNN	17	FLYNN	21959
O'FLYNN	588	FLYNN	21959
O'FARRELL	912	FARRELL	21937
O'DONOVAN	1528	DONOVAN	20505
O'DUFFY	30	DUFFY	20451
MCCAIN	326	CAIN	17078
O'BOYLE	859	BOYLE	16204
O'CALLAGHAN	3130	CALLAGHAN	15492
MCKENNY	462	KENNY	14276
O'MAHONEY	748	MAHONEY	13101
O'DALY	51	DALY	12957
O'REGAN	909	REGAN	12947
O'DRISCOLL	1301	DRISCOLL	12774
CANN	9045	MCCANN	11692
MAHON	5040	MCMAHON	11522
O'FLANAGAN	198	FLANAGAN	11343
KENNA	899	MCKENNA	11255
O'DOHERTY	413	DOHERTY	11219
LOUGHLIN	1959	MCLOUGHLIN	10835
MCEGAN	48	EGAN	10683
MCCAVANAGH	18	CAVANAGH	8535
NALLY	397	MCNALLY	8365
MCMULLEN	3673	MULLEN	8244
O'REILLY	4966	REILLY	8179
O'LEARY	7021	LEARY	8098
O'KANE	693	KANE	7524
MCKAVANAGH	5	KAVANAGH	6693
MCKEATING	252	KEATING	6543
MCCAHILL	78	CAHILL	6541
O'SHEA	5481	SHEA	6340
O'GRADY	3345	GRADY	6310
MCGLYNN	1487	GLYNN	6152
GUINNESS	369	MCGUINNESS	5938
O'DONOGHUE	2188	DONOGHUE	5532
NULTY	442	MCNULTY	5461

N is the number of deaths, 1838-2007, Continued on next page

Surname	N	Updated Surname	N
O'FLAHERTY	935	FLAHERTY	5014
KEOWN	423	MCKEOWN	4906
COY	2924	MCCOY	4816
O'KEEFE	3462	KEEFE	4812
O'ROURKE	3927	ROURKE	4733
O'HANLON	1327	HANLON	4704
MCGARVEY	360	GARVEY	4686
O'TOOLE	2553	TOOLE	4617
MCGROGAN	112	GROGAN	4511
O'HAGAN	1361	HAGAN	4460
GARRY	1368	MCGARRY	4034
MCTIGHE	583	TIGHE	3421
MCMULLIN	793	MULLIN	3375
MALLEY	2124	O'MALLEY	3295
MCCALLAN	188	CALLAN	2779
CARTY	2746	MCCARTY	2762
MCGILLIGAN	85	GILLIGAN	2722
MCSHERRY	410	SHERRY	2714
O'DONOHUE	224	DONOHUE	2680
O'MAHONY	776	MAHONY	2572
KEEFFE	778	O'KEEFFE	2442
CLUSKEY	195	MCCLUSKEY	2441
O'DONOHOE	174	DONOHOE	2057
O'LOUGHLIN	1319	LOUGHLIN	1959
MCCORRY	177	CORRY	1895
SHAUGHNESSY	1547	O'SHAUGHNESSY	1880
MCDADE	683	DADE	1763
O'RIORDAN	540	RIORDAN	1527
MCDEVITT	345	DEVITT	1445
MCMACKIN	48	MACKIN	1397
MCGLENNON	198	GLENNON	1393
HALLORAN	893	O'HALLORAN	1318
MCGEOGHEGAN	62	GEOGHEGAN	1254
MARA	916	O'MARA	1161
O'HANRAHAN	35	HANRAHAN	1128
MCCARROL	33	CARROL	1095
MCCOLGAN	342	COLGAN	1039
CAFFERY	893	MCCAFFERY	999
MEARA	345	O'MEARA	849
KERNAN	504	MCKERNAN	844
MCLAFFERTY	43	LAFFERTY	769
MCGAHAN	462	GAHAN	740
O'BEIRNE	326	BEIRNE	729
MCCREEDY	212	CREEDY	701
CARRON	467	MCCARRON	697
CUSKER	70	MCCUSKER	651
MCCULLY	566	CULLY	598
MCMACKEN	0	MACKEN	507
MCCONVEY	65	CONVEY	487
MCCASHIN	14	CASHIN	472
MCCALVEY	22	CALVEY	458
O'BRYNE	35	BRYNE	417
GREAVY	26	MCGREAVY	381
O'HERLIHY	51	HERLIHY	374
MCTEER	152	TEER	374
KITTRICK	33	MCKITTRICK	372
O'RIELLY	49	RIELLY	350

N is the number of deaths, 1838-2007, Continued on next page

Surname	N	Updated Surname	N
CRUDDEN	57	MCCRUDDEN	346
O'RORKE	213	RORKE	328
MCCOLLUM	180	COLLUM	318
O'HEHIR	43	HEHIR	302
CLOY	64	MCCLOY	294
MCKERNEY	91	KERNEY	290
NERNEY	157	MCNERNEY	272
CUMISKEY	247	MCCUMISKEY	264
QUEENEY	157	MCQUEENEY	244
CUDDEN	83	MCCUDDEN	225
MCCANNY	6	CANNY	192
CARTIN	159	MCCARTIN	192
MCGAVIGAN	8	GAVIGAN	180
MCCOMISKEY	57	COMISKEY	163
MONAGLE	10	MCMONAGLE	157
CLENAGHAN	31	MCCLENAGHAN	139
ANANEY	0	MCANANEY	118
CRICKARD	36	MCCRICKARD	100
O'RAHILLY	21	RAHILLY	97
GRANAGHAN	45	MCGRANAGHAN	84
MCPOLIN	38	POLIN	82
MCLOUGHNEY	16	LOUGHNEY	79
SYOCK	0	SYMCOCK	70
O'CALLAGHAM	0	CALLAGHAM	68
CLARNAN	0	MCCLARNAN	67
CLAFFERTY	4	MCCLAFFERTY	60
NIFFE	0	MCNIFFE	59
ALHONE	0	MCALHONE	56
ELRUE	0	MCELRUE	45
MCTEGGART	10	TEGGART	43
AREAVEY	0	MCAREAVEY	42
ADOREY	0	MCADOREY	39
ILHONE	0	MCILHONE	39
GUONE	0	MCGUONE	35
GURREN	16	MCGURREN	34
GAGHEY	0	MCGAGHEY	33
ENIRY	0	MCENIRY	28
ILHATTON	0	MCILHATTON	23
ILLMURRAY	0	MCILLMURRAY	23
LOUGHIN	0	MCLOUGHIN	21
ELEARNEY	0	MCELEARNEY	17
GENNITY	0	MCGENNITY	17
SHEFFREY	0	MCSHEFFREY	17
ALENEY	0	MCALENEY	16
KEEFRY	0	MCKEEFRY	16
STRAVOCK	0	MCSTRAVOCK	15
ALISKEY	0	MCALISKEY	12
CUSKEY	0	MCCUSKEY	12
MCCAHERTY	0	CAHERTY	11
GAVOCK	0	MCGAVOCK	11
ILMAIL	0	MCILMAIL	11
MANNIMAN	0	MCMANNIMAN	11
MURPHY-CONNOR	0	MURPHY-O'CONNOR	11
ANAW	0	MCANAW	10
PHILOMEY	0	MCPHILOMEY	9
CARTER-GRATH	0	CARTER-MCGRATH	8
ANOY	0	MCANOY	8

N is the number of deaths, 1838-2007, Continued on next page

Surname	N	Updated Surname	N
COY-HILL	0	MCCOY-HILL	8
ILVAR	0	MCILVAR	8
ELHENNY	0	MCELHENNY	7
ERLAINE	0	MCERLAINE	7
CROSBIE-DONNELL	0	CROSBIE-MCDONNELL	6
AVINCHEY	0	MCAVINCHEY	6
CALLISKEY	0	MCCALLISKEY	6
GLEISH	0	MCGLEISH	6
CARROLL-ARDLE	0	CARROLL-MCARDLE	5
ALERNON	0	MCALERNON	5
ASTOCKER	0	MCASTOCKER	5
ATASNEY	0	MCATASNEY	5
NAIR-WILSON	0	MCNAIR-WILSON	5
ATACKNEY	0	MCATACKNEY	4
CAGHY	0	MCCAGHY	4
CUNE-COLBERT	0	MCCUNE-COLBERT	4
ERLEANE	0	MCERLEANE	4
GLEENON	0	MCGLEENON	4
ILKENNY	0	MCILKENNY	4
MENAMAN	0	MCMENAMAN	4
BARRY-CALLAGHAN	0	BARRY-O'CALLAGHAN	3
DILLON-NALLY	0	DILLON-MCNALLY	3
ANEANEY	0	MCANEANEY	3
ANENNY	0	MCANENNY	3
CONIGLEY	0	MCCONIGLEY	3
DOWELL-POLKE	0	MCDOWELL-POLKE	3
GUGGON	0	MCGUGGON	3
KEAGNEY	0	MCKEAGNEY	3
KEEFREY	0	MCKEEFREY	3
KLIZUK	0	KLIMCZUK	2
BRIDE-HARROW	0	MCBRIDE-HARROW	2
CONNELLOGUE	0	MCCONNELLOGUE	2
CUE-SMITH	0	MCCUE-SMITH	2
DERMOTT-PAINE	0	MCDERMOTT-PAINE	2
ELHENNON	0	MCELHENNON	2
ELVANNA	0	MCELVANNA	2
GEOUCH	0	MCGEOUCH	2
GOWAN-SCANLON	0	MCGOWAN-SCANLON	2
INRUE	0	MCINRUE	2
SARSTEDT-CARTHY	0	SARSTEDT-MCCARTHY	2
BINGHAM-GUINNESS	0	BINGHAM-MCGUINNESS	0
FITZPATRICK-GOUGH	0	FITZPATRICK-MCGOUGH	0
HANNAN-DWYER	0	HANNAN-O'DWYER	0
ALARNEY	0	MCALARNEY	0
ALERNEY	0	MCALERNEY	0
ALORAN	0	MCALORAN	0
ANARNEY	0	MCANARNEY	0
ANESPY	0	MCANESPY	0
GUICKIN	0	MCGUICKIN	0
KIVERIGAN	0	MCKIVERIGAN	0

Table 9: Adjusted Irish Stem Names (Mc and O')

Wealth Regressions

To investigate whether the "Irish" effect on probated wealth is robust when controlling for age at death, we use the linked PPR-Death data to estimate two models. First we look at the extensive

margin, the probability of probate (of achieving 'significant' wealth at death).

$$Prob(p_i) = \alpha + D_i^F + Age_i + Age_i^2 + \sum D^E \tag{6}$$

where p_i is a categorical variable indicating whether an individual i was probated, α is a constant, D^F is a categorical variable code to one where an individual i has a typically female first name, Age is age at death, and D^E are categorical variables indicating ethnicity of an individual's surname. The results of this regression are reported in table 10. Table 11 controls for distrcit of death.

Table 10: Probability Probated and Ethncity, Linked Data: Deaths->PPR, controlling for Age at Death

	Probated (1/0)*100					
	1866-1899		1900-49		1950-1992	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	−6.12***	−6.20***	−5.42***	−5.64***	−2.11***	−.88***
	(.03)	(.03)	(.03)	(.03)	(.02)	(.02)
Welsh	.54***	.62***	2.80***	3.44***	−.64***	−.85***
	(.11)	(.11)	(.10)	(.10)	(.05)	(.05)
Scottish	1.25***	1.61***	−.60***	−.39***	−2.89***	−3.01***
	(.13)	(.13)	(.11)	(.11)	(.06)	(.06)
Irish	−6.99***	−6.28***	−11.14***	−10.23***	−5.53***	−6.09***
	(.10)	(.10)	(.10)	(.09)	(.05)	(.05)
Other	−1.46***	−.88***	−5.29***	−4.82***	−2.45***	−2.45***
	(.18)	(.18)	(.12)	(.12)	(.06)	(.06)
Age at Death Quadratic?		✓		✓		✓
Observations	3,155,398	3,155,398	7,716,988	7,716,988	14,274,538	14,274,538
R^2	.01	.02	.01	.02	.002	.01

Note: *p<0.05; **p<0.01; ***p<0.001
Linear Probability Model (OLS), English is the omitted Group.

Table 12 reports the results of the regression

$$log(w_i) = \alpha + D_i^F + Age_i + Age_i^2 + \sum D^E \tag{7}$$

where w_i is *probated* real wealth. Table 13 controls for county of death.

Table 11: Probability Probated and Ethncity, Linked Data: Deaths->PPR, controlling for Age at Death and District of Death

	Probated (1/0)*100					
	1866-1899		1900-49		1950-1992	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	−6.20***	−6.33***	−5.64***	−5.85***	−.88***	−1.23***
	(.03)	(.03)	(.03)	(.03)	(.02)	(.02)
Welsh	.62***	.40**	3.44***	3.09***	−.85***	−.19***
	(.11)	(.13)	(.10)	(.11)	(.05)	(.05)
Scottish	1.61***	1.53***	−.39***	−.40***	−3.01***	−2.75***
	(.13)	(.13)	(.11)	(.11)	(.06)	(.05)
Irish	−6.28***	−5.59***	−10.23***	−9.16***	−6.09***	−5.09***
	(.10)	(.11)	(.09)	(.09)	(.05)	(.04)
Other	−.88***	−.97***	−4.82***	−3.78***	−2.45***	−2.99***
	(.18)	(.18)	(.12)	(.12)	(.06)	(.06)
Age at Death Quadratic?	✓	✓	✓	✓	✓	✓
District Fixed Effects?		✓		✓		✓
Observations	3,155,398	3,155,398	7,716,988	7,716,988	14,274,538	14,274,538
R^2	.02	.03	.02	.04	.01	.08

Note: *p<0.05; **p<0.01; ***p<0.001
Linear Probability Model (OLS), English is the omitted Group.

Table 12: Probated Wealth and Ethncity, controlling for Age at Death

	log(Real Wealth)					
	1866-1899		1900-49		1950-2007	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	−.119***	−.163***	−.189***	−.275***	−.181***	−.238***
	(.003)	(.006)	(.001)	(.002)	(.001)	(.002)
Welsh	−.265***	−.125***	−.145***	−.038***	−.043***	.019***
	(.006)	(.016)	(.002)	(.006)	(.003)	(.005)
Scottish	.254***	.295***	.194***	.224***	.157***	.171***
	(.008)	(.012)	(.003)	(.004)	(.003)	(.004)
Irish	−.009	.097***	−.145***	−.087***	−.218***	−.181***
	(.011)	(.018)	(.004)	(.006)	(.004)	(.005)
Other	.490***	.616***	.237***	.339***	.307***	.326***
	(.014)	(.022)	(.005)	(.007)	(.006)	(.007)
Age at Death Quadratic?		✓		✓		✓
Observations	1,004,139	345,756	4,691,333	2,146,999	3,703,560	1,941,926
R^2	.006	.013	.006	.022	.007	.013

Note: *p<0.05; **p<0.01; ***p<0.001
OLS, English is the omitted Group.

Table 13: Probated Wealth and Ethncity, controlling for Age at Death and County

	log(Real Wealth)					
	OLS	*felm*	*OLS*	*felm*	*OLS*	*felm*
	1866-1899		1900-49		1950-2007	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	$-.163^{***}$	$-.278^{***}$	$-.275^{***}$	$-.287^{***}$	$-.238^{***}$	$-.245^{***}$
	(.006)	(.010)	(.002)	(.002)	(.002)	(.002)
Welsh	$-.125^{***}$	.015	$-.038^{***}$	$.041^{***}$	$.019^{***}$	$.056^{***}$
	(.016)	(.031)	(.006)	(.006)	(.005)	(.006)
Scottish	$.295^{***}$	$.345^{***}$	$.224^{***}$	$.251^{***}$	$.171^{***}$	$.198^{***}$
	(.012)	(.021)	(.004)	(.004)	(.004)	(.004)
Irish	$.097^{***}$	$.066^{*}$	$-.087^{***}$	$-.065^{***}$	$-.181^{***}$	$-.149^{***}$
	(.018)	(.031)	(.006)	(.006)	(.005)	(.006)
Other	$.616^{***}$	$.583^{***}$	$.339^{***}$	$.306^{***}$	$.326^{***}$	$.333^{***}$
	(.022)	(.036)	(.007)	(.007)	(.007)	(.007)
Age at Death Quadratic?	✓	✓	✓	✓	✓	✓
County Fixed Effects?		✓		✓		✓
Observations	345,756	122,704	2,146,999	2,047,462	1,941,926	1,865,598
R^2	.013	.043	.022	.032	.013	.025

Note: *p<0.05; **p<0.01; ***p<0.001
OLS, English is the omitted Group.

Results using Different Thresholds for Assigning Ethnicity from the 1911 Census

Our favored measure of ethnicity is assigned for an individual surname based on the distribution of place of birth of the holders that surname in the 1911 census of England and Wales. Using a two step rule, we first attribute to each surname an ethnic origin based upon the most frequent country of birth. Where there is a country other than England or Wales which accounts for 20% or over of the births of that surname, we update the ethnic origin to that country. This procedure works to attribute correctly all of the example surnames in table 2).

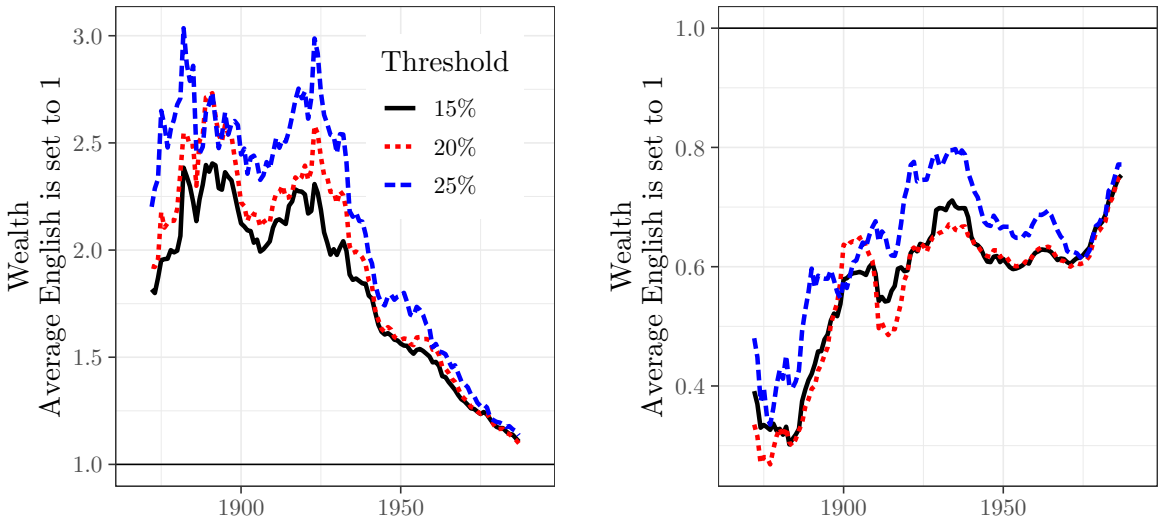
Here we compare our average wealth results based upon the 20% threshold used in the paper with assignments based on 15% and 25% thresholds. Figure 14 reports the average wealth for the Irish and Scottish ethnic groups for these varying thresholds. We exclusively focus on these two groups as they are the major 19th century migrant groups to England and Wales, and thus the threshold dynamics are expected to be similar.

For the Irish the 15% and 20% cutoffs produce almost identical results with some modest divergence in the 1860s, and early 20th century. However both the Irish and the Scottish, a 25% threshold produces higher estimates of wealth for most of the sample period. A 15% rule also produces slightly lower estimates of wealth for the Scottish.

There are 506,441 "Irish" with 19,868 surnames assigned using a 15% cut-off, 360,912 with 18,316 surnames using a 20% cut-off, and 206,091 with 17,201 surnames using a 25% cutoff. Thus the counts per surname decline as the assignment threshold rises, from 25.49 per surname using a 15% rule, to 19.7 for a 20% rule, and 11.98 for a 25% rule. Therefore employing the higher cut-off has the undesired effect of selecting for more unusual, smaller names. This is potentially problematic if rarer names themselves are correlated with wealth.

Figure 15 demonstrates that for the English, rarer names are indeed correlated with higher wealth. This figure reports wealth 1858-1992, stratified for English surnames by frequency range in the universe of deaths, 1838-1900. Thus utilizing a higher selection cutoff in assigning ethnicity may result in an unrepresentative sample as it selects for rarer names, which are powerfully correlated with wealth.

Thus, as a sanity check we calculated average wealth by year for the top 15 Irish surnames reported previously in table 3. We do this as these top 15 Irish surnames are commonly accepted as being "Irish". Figure 16 reports the average wealth for this group of commonly accepted Irish aligns almost perfectly with that generated from our 15% and 20% threshold cutoffs. In the next section we compare our estimates with those using an entirely different method of attributing surnames to ethnicity, and find our cutoff to be broadly in line with that too. Thus we interpret our 15% and 20% cutoffs for defining ethnicity as representing the average experience of the Irish in England, with the 25% representing a positively selected right tail of the Irish in England status distribution.



(a) Scottish

(b) Irish

Figure 14: Comparing Average Wealth, 1858-1992, using Varying Thresholds for a Surname's Ethnicity

Note: We assign ethnicity using the individual records of the 1911 census of England and Wales. For each surname we inspect the distribution of birth of the holders of that surname. Using a two step rule, we first attribute to each surname an ethnic origin based upon the most frequent country of birth. Where there is a country other than England or Wales which accounts for 20% or over of the births of that surname, we update the ethnic origin to that country. For robustness we here compare that 20% rule with a 15% rule, and a 25% rule.

Figure 15: The Variation in Average Wealth by the 1858-1900 Frequency of a Surname, 1858-1992

Notes: Range refers to Frequency range of a surname in the universe of death records, 1838-1900. The figure illustrates that rarer surnames are richer than common surnames. 'Common' English are surnames in 10,000 and over range.

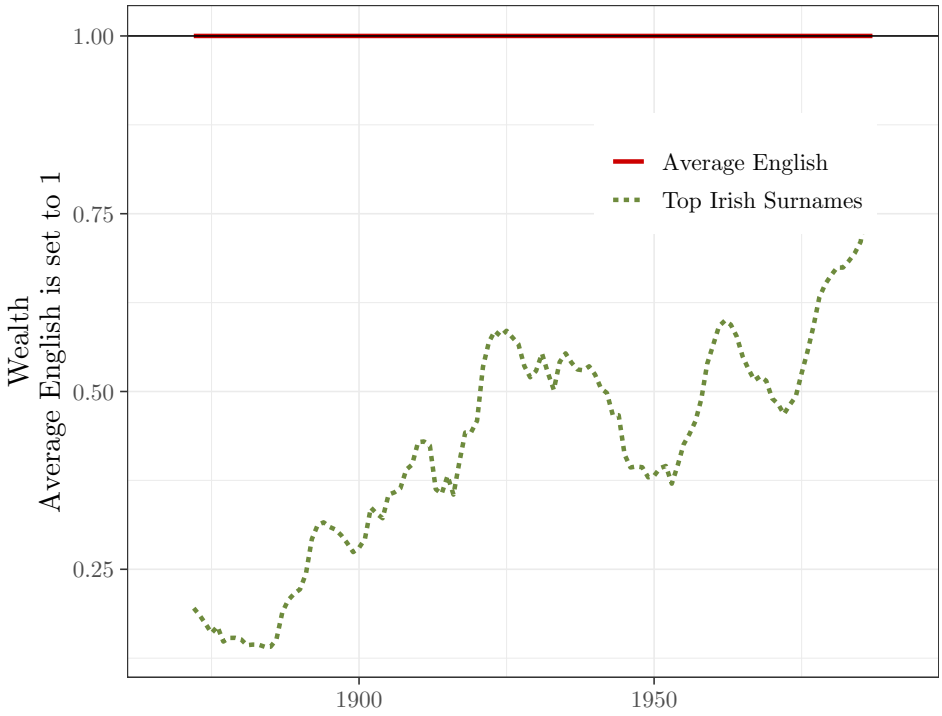


Figure 16: Wealth for the Average English compared with those holding one of the Top 15 Irish Surnames
Notes: We took the top 15 Irish surnames as listed at https://forebears.io/ireland/surnames, a website that has aggregated a considerable volume of data on contemporary global surname distributions.

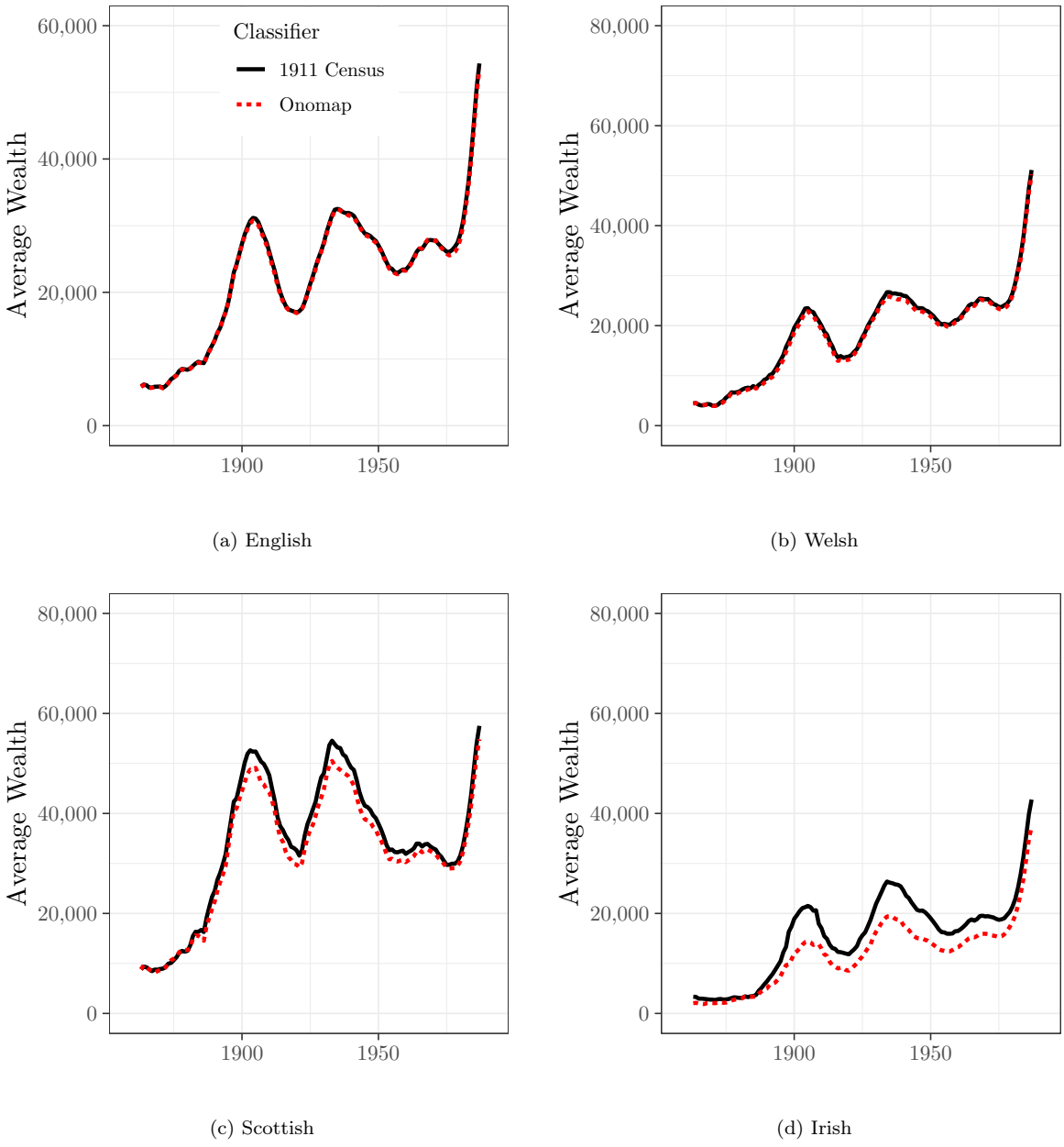
Results using an Alternative Ethnic Classification

How robust are these patterns to a entirely different method of ethnic classification? We compare the results of our 1911 ethnicity assignment with that of 'Onomap', a classification system developed by Paul Longley and numerous collaborators at University College London. Using billions of contemporary records from telephone directories and electoral registers, from nearly all countries in the World, a network analysis clusters surnames together based upon shared forenames (Mateos et al. (2011)). These clusters map on to known ethnocultural groups. An example for the Irish would be an observed cluster containing surnames such as *Murphy*, *McCarthy*, *Kelly*, and *O'Shea*, linked to each other through shared, distinctively Irish, forenames such as *Cormac*, *Bridget*, *Niall* and *Sorcha*.

Figure 17 reports the average wealth for the British and Irish ethnic groups for both classifications. They are identical for the English, the Scottish, and the Welsh. However, for the Irish the trends are different. The Onomap classifier results in wealth estimates substantially lower than that of the 1911 census assignment used here.

Figure 18 reports the infant mortality rate for the British and Irish ethnic groups for both classifications. The different classifications produce identical results.

One possible explanation for the divergence between the methods for Irish Wealth is the *contemporary* nature of the Onomap classifier. Over time, successful Irish could integrate into the English and adopt English forenames for their children. These Irish would then be classified as 'English' by Onomap. Of course it also could be that Onomap better classifies ethnicity than our 1911 Census classifier. (Or vice versa of course.) The 1911 census based classifier still results in Irish wealth significantly below that of the English. It may be that our choice of classifier is an overestimate of Irish wealth, relative to Onomap, and thus an *underestimate* of the true Irish-English, in England, wealth gap. The conclusions from the main analysis are unchanged.



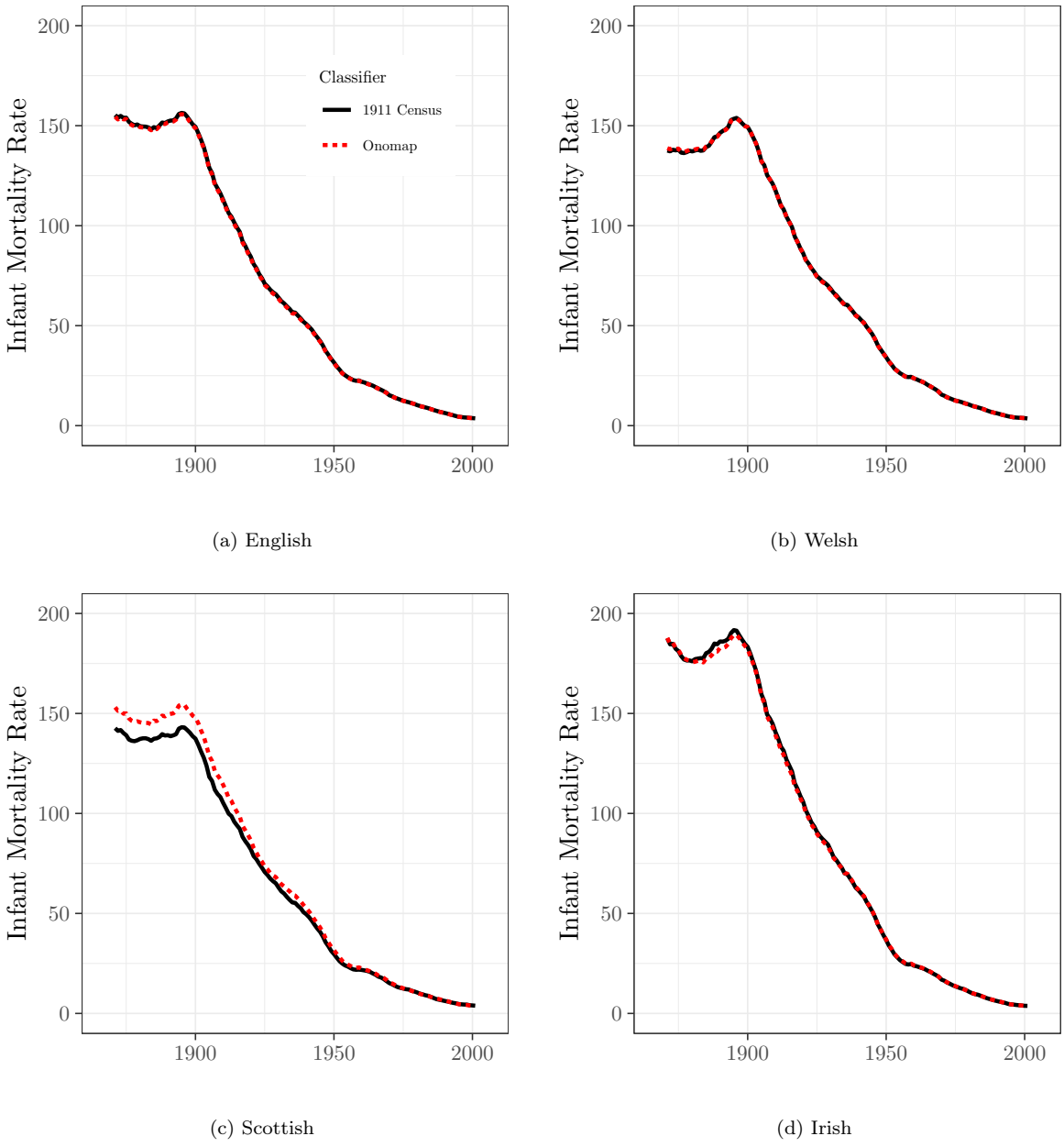
(a) English

(b) Welsh

(c) Scottish

(d) Irish

Figure 17: Comparing Average Wealth, 1858-1992, using Alternative Ethnic Classifier



(a) English

(b) Welsh

(c) Scottish

(d) Irish

Figure 18: Comparing Infant Mortality Rates, 1866-2007, using Alternative Ethnic Classifier

Extra Tables and Figures

Table 14: Summary Statistics, Wealth

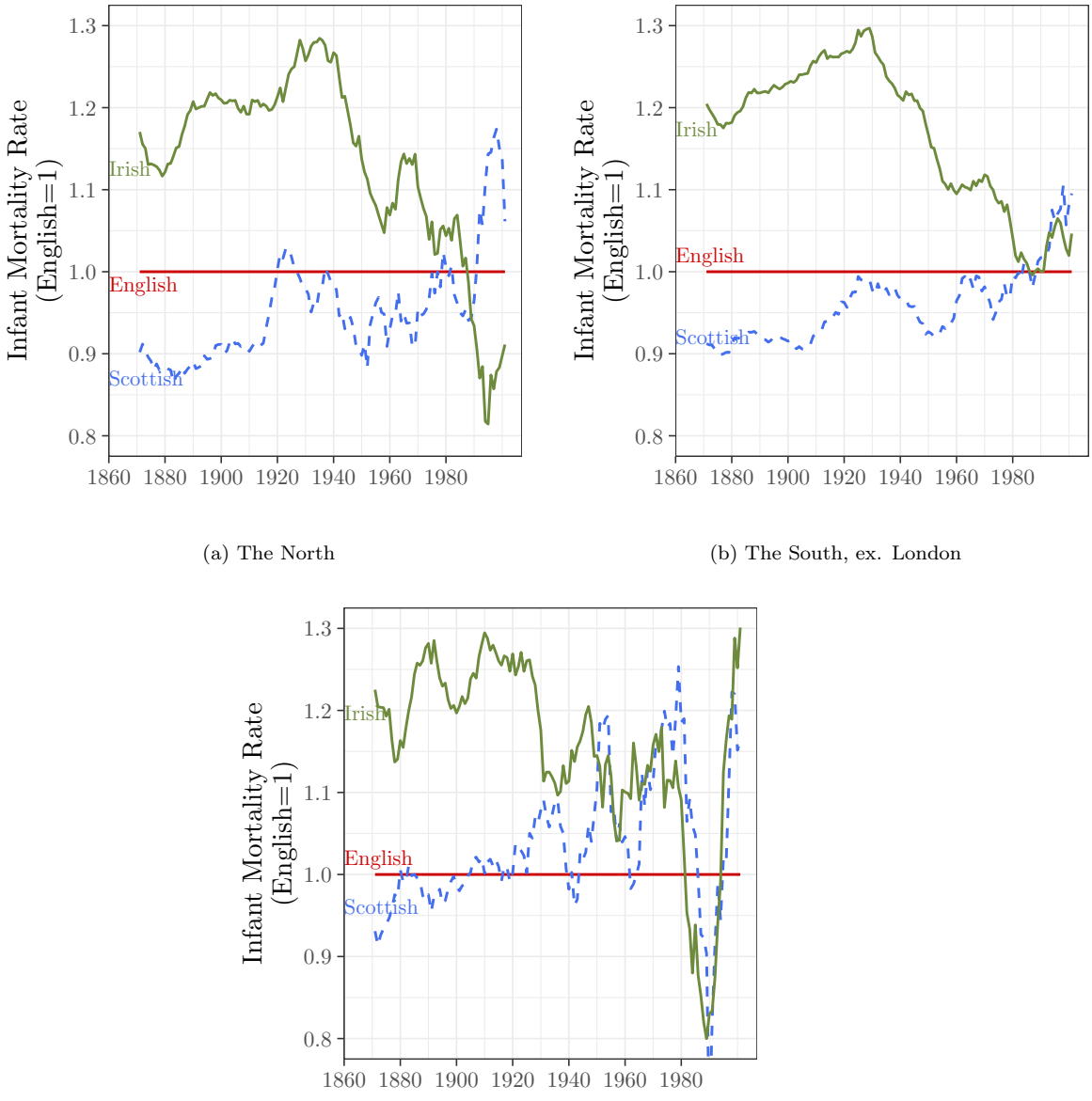
	N			Real Wealth				
Ethnicity	Deaths	Probates	Probate Rate	Min.	Max.	Mean	Median	Top 1% Prop.
All	70,045,750	15,246,247	0.22	0	1,257,371,575	23,856	895	0.010
English	61,354,562	13,379,209	0.22	0	1,257,371,575	23,967	895	0.010
Welsh	5,700,027	1,225,467	0.21	0	165,769,426	19,543	895	0.008
Scottish	813,189	219,782	0.27	0	113,345,260	43,683	1,204	0.020
Irish	1,648,773	258,650	0.16	0	105,187,746	14,932	849	0.005

Source: Universe of Deaths 1837-2007 and Probates 1858-1992.

Table 15: Summary Statistics, Probate Rate, 1996-2018

	N		
Ethnicity	Deaths	Probates	Probate Rate
All	11,739,085	5,601,612	0.48
English	10,165,342	4,874,073	0.48
Welsh	942,959	454,052	0.48
Scottish	216,224	104,944	0.49
Irish	414,560	168,543	0.41

Source: Database of the number of deaths and the number of probates, by surname, 1996-2018.



(a) The North

(b) The South, ex. London

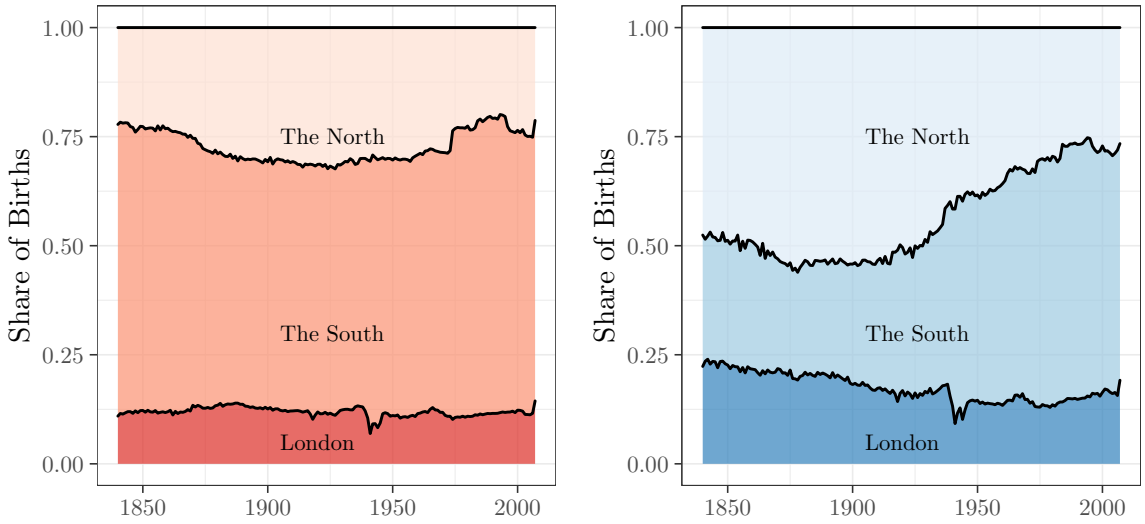
(c) London

Figure 19: Regional Differences in the Relative Infant Mortality of the Irish
Note: The English baseline is established by region.

Table 16: Summary Statistics, Infant Mortality, 1866-2006

Ethnicity	1866-1929			1930-2006		
	Births	Infant Deaths	IMR	Births	Infant Deaths	IMR
All	54,062,848	7,035,768	130.1	56,449,908	1,230,969	21.8
English	47,024,384	6,125,985	130.3	44,359,607	1,002,965	22.6
Welsh	4,447,423	557,552	125.4	4,159,910	99,061	23.8
Scottish	515,531	60,800	117.9	970,571	18,061	18.6
Irish	1,052,308	163,358	155.2	2,214,129	47,564	21.5

Source: Universe of Births and Deaths, 1866-2006. IMR is the infant mortality rate, per 1,000 births.



(a) The Welsh

(b) The Scottish

Figure 20: The Regional Distribution of Births, the Welsh and the Scottish